ネットワーク・デザイン問題とそのグラフ構造 (決定理論と最適化アルゴリズム)

毛利, 裕昭

数理解析研究所講究録 1409 巻 2005 年 195-204

ネットワーク・デザイン問題とそのグラフ構造

早稲田大学 商学部，理工学部 数理科学研究所
毛利 裕昭 (Hiroaki Mohri)
School of Commerce, Waseda Institute of Mathematics
Waseda University.
mohri@waseda.jp, mohri@member.ams.org

1 はじめに

ネットワーク・デザイン問題は，多品種流問題と各アークに関する設備配置を同時に考える問題である．多種多様なバリエーションをこの問題は持っている．応用は，計算機ネットワーク，通信ネットワーク，交通ネットワークの計画等にわたる．特にインターネット社会ではその重要性は増している．

ここでは，あるグラフ構造上でのネットワーク・デザイン問題を論じる．こうした離散最適化問題を論じる場合には，完全グラフ上で考えるのが一般的である．特殊なグラフ上での最適化問題は，提携関数値の値が離散最適化問題の値として研究される協力ゲームなどにみられる．その場合，対象とする問題のグラフの構造を Under Lying Graph と呼ぶ．この研究における動機は，「離散最適化ゲーム」における Under Lying Graph の構造によって協力ゲームにおける求解アルゴリズムが変化しうるということである．(協力ゲームの解は，多種多様であるがここでは特に述べない.) つまり，アルゴリズムをグラフの構造を生かしたものにすれば，そのアルゴリズムが計算の複雑性理論においてどのようなクラスに属するかも当然変化することになる．ここでは，ゲーム理論への具体的応用までは述べない．

上記で「グラフ構造」という用語を特に何も定義することなくも一般用語として用いている．しかし，離散最適化問題を考える上では，体系的に論じることが望ましい．グラフ構造に関して，トポロジカルなグラフ理論的考察とグラフのアークとノードに関する属性に関する考察が必要になってくる．このことについて簡単な試案を述べる．

その上で，一般には，$\mathcal{NP}$ 困難な問題であるネットワーク・デザイン問題に対して，木の上および直並列グラフの特殊なクラス上で考えた場合の結果を述べる．これらは，木や直並列グラフの特性を生かすことにより解を強多項式時間で得ることができることを示す．特に直並列グラフ関係の場合は，直並列グラフを生成する（分解する）アルゴリズムの特徴を利用する．

2 グラフの構造とは

グラフの構造は，そのトポロジカルな性質を中心に探求されてきたとと思われる．まず，グラフ理論に登場するグラフの構造は多岐にわたり，完全グラフ，連結グラフ，木，森サイクル，単純グラフ（並列枝，自己閉路なし），並列グラフ，二部グラフ，平面グラフ，k-正則グラフなどがある．これらは，体系的に考えられたものではなく研究テーマに応じて生み出された（発見された）ものであると想像される．

最適化ということが前提にあるとマトロイド等の観点からは，木に近いか等が一つの基準になりそうである．木や森では，品種（ノードペア）が与えられると一意に道（有向道）が決定できる．この性質はネットワークの最適化では良い性質の一つといえる．しかし，それもアイディアの一つにすぎないであろう．

上記に関連していることとして品種の通りうる道に関する研究があろう．グラフ理論で古典的な定理である Menger の定理や連結度，リンク度に関わる研究が援用されうる可能性がる．

また，Diestel[1] の「5.5 理想グラフ（Perfect Graph）」でのコメントが興味深い．彼は Lovasz の研究とサーベイを引用して，理想グラフの最適化問題での重要性を述べている．この件に関しては多種多様な意見があろう．

いずれにせよ，最適化問題を前提としてグラフの構造および情報というものを考える時ノードとアークの各々の集合に注目した $G(N,A)$ という情報だけを使うのでないことが多い．では，どのような情報を使うのかといえば，各ノード，各アークについての属性というべきものとなろう．以下，必要と考える情報を列挙する．

$W(N)$：ノード上の重み（ex. 施設の費用）
$W(A)$：アーク上の重み（ex. 枝敷設費用）
$Q(N)$：ノードの容量 (ex. 施設の容量)
$Q(A)$：アークの容量（ex. 枝の容量）
$C(N)$：ノードに入れるもの 1 単位の費用
$C(A)$：アーク上で 1 単位のフローを流す為の費用

これらから，$G(N,A:W(N),W(A),Q(N),Q(A),C(N),C(A))$ を考えれば全てのとはいえなくともかなり一般的な離散最適化問題のグラフ構造を表現可能ではないかという気がする．本論であつかうネットワーク・デザイン問題に関しては，$G(N,A:W(N),W(A),Q(A),C(A))$ であるが，これに施設配置問題を絡ませると先にのべたすべての情報が必要となる．

3 ネットワーク・デザイン問題の定式化

記号

N	ノード集合 $\{1,\ldots,n\}$
A	アーク集合
K	品種の集合（ノードペアの集合），$K \subseteq N \times N$
$O(k)$	品種 $k \in K$ の出発ノード
$D(k)$	品種 $k \in K$ の終点ノード
F_{ij}	アーク (i,j) が，敷設された時の固定費用
c_{ij}^k	品種 k 1単位が，アーク (i,j) を通貨する時のフロー費用

以下のように記述されるネットワーク・デザイン問題を **Basic Network Design Problem** とし，**BNDP** と略記する．

BNDP:

$$\min \sum_{(i,j)\in A} \sum_{k\in K} c_{ij}^k f_{ij}^k + \sum_{(i,j)\in A} F_{ij} y_{ij} \tag{1}$$

subject to

$$\sum_{j\in N-\{i\}} f_{ij}^k - \sum_{j\in N-\{i\}} f_{ji}^k = \begin{cases} 1 & i = O(k) \\ 0 & \forall i \in N - \{O(k), D(k)\} \\ -1 & i = D(k) \end{cases} \quad \forall k \in K \tag{2}$$

$$\sum_{k\in K} f_{ij}^k \leq |K| y_{ij} \quad \forall (i,j) \in A \tag{3}$$

$$f_{ij}^k \geq 0 \quad \forall (i,j) \in A, \forall k \in K \tag{4}$$

$$y_{ij} \in \{0,1\} \quad \forall (i,j) \in A \tag{5}$$

後での議論の為に上記の定式化では並列枝が無視されていることに注目しておく必要がある．

この定式化において，目的関数 (1) は，固定費用（アーク敷設コスト）とフロー費用の和をこのネットワーク上で最小化している．

制約条件 (2) は，フロー保存式である．(3) は，アーク (i,j) が敷設されていなければ，そのアークを通過するフロー量は 0 でなくてはならないことを示す．

BNDP は，Garey [3] で完全グラフ上で $\mathcal{NP}$ 困難であることが証明されている．その為に大規模な BNDP は，厳密解を得るのが計算量の複雑性理論の上で難しい．もし BDNP を通信や交通の現実問題に応用しようとすると，多項式時間で解けないので様々な近似解法が提案されている．

4 多項式時間で解けるネットワーク・デザイン問題

4.1 完全グラフ上の最小全域木問題

マトロイド理論などにおいて，最小全域木問題は最も重要な問題である．この問題はクラスカルの貪欲アルゴリズムで完全グラフ（条件を弱めて，単に連結グラフでも良い）上で多項式時間で解くことができる．**BNDP** に関しては，以下のように上記で述べた定式化でおくことによって，最小全域木問題はネットワーク・デザイン問題の特殊なケースであることがわかる．ここでは，ノード 1 をルート・ノードと見なしている．

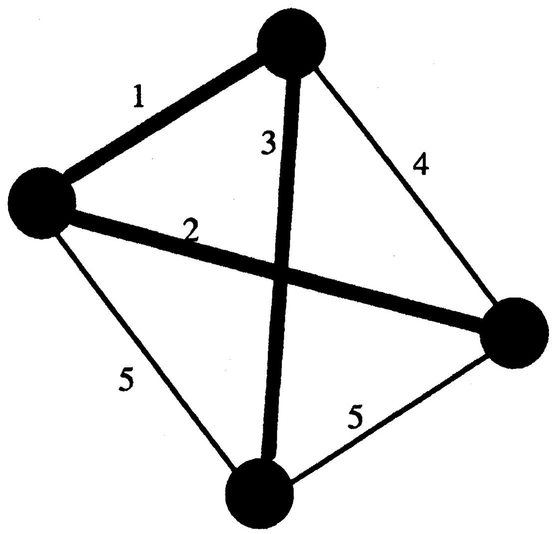

図 1: 最小全域木の例

$\{1\} = \{O(k) | k \in K\}$ また，目的関数において $d^k_{ij} = 0$ とおき，以下の定式化を得る．

$$\min \sum_{(i,j)\in A} F_{ij} y_{ij} \tag{6}$$

4.2 木もしくは森に関するネットワーク・デザイン問題

森においては，品種 $k \in K$ に対して $O(k)$ と $D(k)$ が同じ部分グラフ（つまり木）上にあると場合のみを考える．よって，森が木である場合を証明すればそれで十分である．この状況において，それぞれの品種に対して，ネットワーク・フローのパスは非常に容易に決定できる．というのは，木に対しては，出発点から終点までの道はその最短経路で一意に決定できるからである．

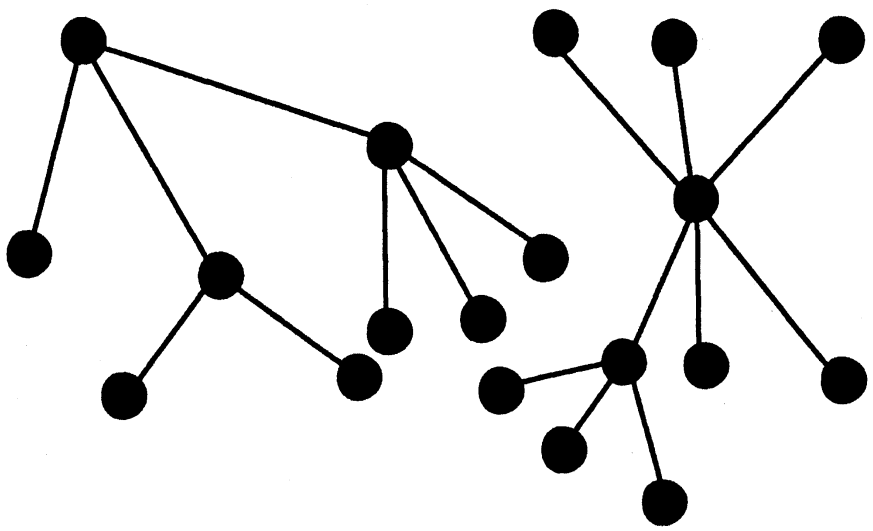

図 2: 森の例

4.3 タンデム並列枝グラフに対するネットワーク・デザイン問題

タンデム並列枝グラフは，一般的なグラフ理論や離散数学の用語ではない．**直並列グラフ**の部分集合である．**並列枝グラフ**，**タンデム**の定義は順に述べる．

4.3.1　並列枝グラフ

まず，**並列枝グラフ**は，ここでは並列枝を持つグラフと定義する．　**単純並列枝グラフ**は，ノードが２つの**並列枝グラフ**とする．

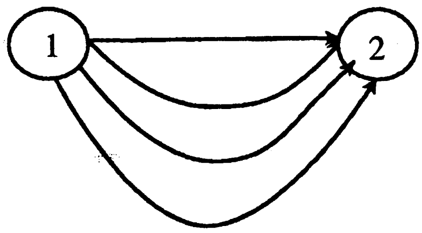

図 3: 単純並列枝グラフの例

ここでは　**単純並列枝グラフ**の場合を取り上げる．２つのノード，ノード１，ノード２の間に並列枝が存在する．図３では，その様なグラフの例である．このグラフに対応する現実現象は多々見られる．例えば，光ファイバー・ケーブルをノード１とノード２の間に敷設する際に一本あたりの光ファイバー・ケーブルの容量に制約があり，複数のケーブルを２つのノード間に敷設することは決してめずらしいことではない．ここで各光ファイバー・ケーブルは固定費用とフローコストを持つ．追加記号を以下に示す．

d：　ノード１からノード２への品種の需要
q：　各枝の容量
p：　$p = [d/q]$
r：　$d = r\ (mod\ q),\ 0 \leq r < q$
n：　図３における並列枝の数

以下を仮定する

$$(\min\{F_e | e \in A\})/q > \max\{c_e | e \in A\}. \tag{7}$$

この仮定は非常に自然なものである．それは，一般にフロー一単位あたりの固定コストは（一単位の）フローコストよりも大きい．

この問題の解は，以下のアルゴリズムで得ることができる．

単純並列枝グラフに対するBNDPの最適化アルゴリズム

(Step 1) $\{F_e + c_e q | e \in A\}$ をソートする．ここで $F_{e_1} + c_{e_1} q \leq F_{e_2} + c_{e_2} q, \ldots, F_{e_n} + c_{e_n} q$ である．また，$A_F = \{e_1, e_2, \ldots, e_p\}$ と定義しておく．

(Step 2) $\{F_e + c_e r | e \in A\}$ をソートする．ここで $F_{e_1'} + c_{e_1'} r \leq F_{e_2'} + c_{e_2'} r, \ldots, F_{e_n'} + c_{e_n'} r$ である．また，$A_C = \{e_1', e_2', \ldots, e_p'\}$ と定義する．

(Step 3) もし，$A_F \cap A_C = \emptyset$ ならば，A_F の全ての要素は，フローの値を q とする．そして，$e_1{}' \in A_C$ は，フローの値を r とする．そうでなければ，**(Step 4)** へ進む．

(Step 4) 以下のような $k \in A_F$ and $l \in A \backslash A_F$ を探す．それは $\{((F_k + c_k q) + (F_l + c_l r)) - ((F_l + c_l q) + (F_k + c_k r))\} = \{(c_k(q-r) - c_l(q-r)\}$ を最大化するものである．(k^*, l^*) は，最大化ペアを示すものとする．$A_F \backslash \{k^*\} \cup \{l^*\}$ の全ての要素は，フローの値 q とし，k^* はフローの値を r とする．

このアルゴリズムは，ソートを２度と線形探索を用いるだけである．結果，解を強多項式時間で得ることができる．

4.3.2 タンデム並列枝グラフ

図４のような**タンデム並列枝グラフ**を考える.「タンデム」は，待ち行列理論にみられる「タンデム（直列）」と同じである．ここで考えているのは,「**単純並列枝グラフ**を直列にならべたものある．この例では，３つの品種があり，それらはノード１ → ノード２,ノード１ → ノード３,ノード２ → ノード３．各需要は d_1, d_2, d_3 とする．この問題は，上記の**単純並列枝グラフ**の問題に帰着できる．１つめの問題は，ノード１，２間で需要を $d = d1 + d2$ とする．一方，ノード２とノード３の間の需要を $d = d2 + d3$ として問題を解くことである．こうした分解によって**タンデム並列枝グラフ**に対するネットワーク・デザイン問題が強多項式時間で解けることが明らかである．

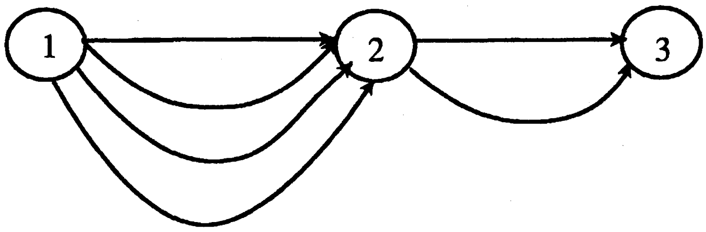

図 4: タンデム並列枝グラフ

5 施設配置を考慮したネットワーク・デザイン問題に関する考察

5.1 施設配置を考慮したネットワーク・デザイン問題とは

ここでは，筆者が毛利 [6] で提案したものを紹介しそれについていくつかの簡単な考察を試みることにする．

<集合等>

N	ノード集合
A	アーク集合
$G(N,A)$	グラフ
M	施設候補集合　$(M \subseteq N)$
K	品種集合, $K \subseteq N \times N$

<定数>

d_{ij}^k　アーク (i,j) に品種 k が 1 単位流される時のルーティング（フロー）費用
F_{ij}　アーク (i,j) のデザイン（敷設）費用
L_i　施設 i が，作られたときの固定費用
l　ネットワーク上で敷設すべき施設の数（所与とする）

<決定変数>

f_{ij}^k　アーク (i,j) における品種 k のフロー量変数
y_{ij}　アーク (i,j) のデザイン（敷設）変数
z_i　施設 i の設置変数

ここでは，解釈のし易い**アークフロー定式化**を記述する．

$$min \sum_{(i,j)\in A}\sum_{k\in K} d_{ij}^k f_{ij}^k + \sum_{(i,j)\in A} F_{ij}y_{ij} + \sum_{i\in M} L_i z_i \tag{8}$$

subject to

$$\sum_{j\in N-\{i\}} f_{ij}^k - \sum_{j\in N-\{i\}} f_{ji}^k = \begin{cases} 1 & i = O(k) \\ 0 & \forall i \in N - \{O(k), D(k)\} \quad \forall k \in K \\ -1 & i = D(k) \end{cases} \tag{9}$$

$$\sum_{k\in K} f_{ij}^k \leq |K| y_{ij} \quad \forall (i,j) \in A \tag{10}$$

$$\sum_{i\in M} z_i = l \tag{11}$$

$$f_{ij}^k \geq 0 \quad \forall (i,j) \in A, \forall k \in K \tag{12}$$

$$y_{ij} \leq z_i \quad \forall i \in M \quad \forall j \in N \tag{13}$$

$$y_{ji} \leq z_i \quad \forall i \in M \quad \forall j \in N \tag{14}$$

$$y_{ij} \in \{0,1\} \quad \forall (i,j) \in A \tag{15}$$

$$z_i \in \{0,1\} \quad \forall i \in M \tag{16}$$

(9) は，ネットワーク問題で典型的に現れるフロー保存則の制約式群である．(10) は，アーク上のフロー量が，使用するアークの容量以下であること示す．(11) は，施設数制約である．(13),(14) は，アークと施設の関係制約である．

5.2 木に対する施設配置を考慮したネットワーク・デザイン問題

木であれば，先に考察をおこなったように品種に対するフローの流れが一意に決定される．したがって，どこにどのように流すか最初に決定してしまう．したがって y_{ij} をどうするべきかはすぐに決定されてしまう．施設の容量制約がこの問題にはないので z_i に関しては L_i が小さいものから greedy に l 個が 1 になる．

5.3 さらに一般的な施設配置を考慮したネットワーク・デザイン問題

上記では，2 章で提案した $G(N,A:W(N),W(A),Q(N),Q(A),C(N),C(A))$ をすべて使用しているわけではない．施設に対して $Q(N),C(N)$ の情報を使っていない．まず，各ノード i での**境界**（すなわちフローの入出力差）$\partial\varphi(i)$ を以下のように定義する．

$$\partial\varphi(i) = \sum_{k\in K}\Big(\sum_{j\in N-\{i\}} f_{ij}^k - \sum_{j\in N-\{i\}} f_{ji}^k\Big) \tag{17}$$

<追加定数>

B_i　施設 i における境界一単位当たりの費用

Q_i　施設 i における容量

ここで，目的関数に

$$\sum_{i\in N} B_i\partial\varphi(i) \tag{18}$$

を加えるとする．

制約式には

$$\partial\varphi(i) \leq Q_i \tag{19}$$

を想定する．もし，木構造の上でこの問題を考えるなら上記の単なるネットワーク・デザイン問題と全くまったく同じ状況になる．制約式が破られていなければ，まったく同じ議論で $L_i + B_i\partial\varphi(i)^*$（ただし，$\partial\varphi(i)^*$ は，フローから計算される境界値）が小さいものから greedy に l 個とればよいことになる．しかし，タンデム並列枝グラフに関してはネットワーク・デザイン問題と同じ議論はできない．ノードと枝の相互関係がきれいに扱えなくなるからである．

6 結果と今後の課題

ネットワーク・デザイン問題とその派生問題である「施設配置を考慮したネットワーク・デザイン問題」に対してグラフの構造が非常に単純なものに関する考察をおこなった．木構造に対する結果はほとんど自明であり確認をしただけといえる程度のものある．今後の課題として，直並列枝グラフのさらに条件を一般化したものについて多項式時間で解けるものがあるかをまず第一にさぐりその上で「施設配置を考慮したネットワーク・デザイン問題」に対してどのようになるかを考えてみたい．また，他の最適化問題についても様々なグラフに対してチェックを行なう．

謝辞

本研究は早稲田大学特定課題研究助成費 2004A-122 の成果の一部である．

参考文献

[1] R. Diestel, Graph Theory 2nd Edition, (Springer-Verlag, 2000).

[2] Y. Mizoguchi, Shortest path length calculation using graph transformations, in Proceedings 6th Joint Conference on Information Sciences, North Carolina, 358-361(2002).

[3] M. R. Garey and D. S. Johnson, A Guide to the Theory of NP-Completeness, (Freeman, 1979).

[4] G. L. Nemhauser and L. A. Wolsey, Interger and Combinatorial Optimization, (Wiley, 1988).

[5] C. Wynants, Network Synthesis Problems, (Kluwer, 2000).

[6] 毛利裕昭, 施設配置を考慮したネットワーク・デザイン問題について, 京都大学数理解析研究所考究録,1297, 89-95(2002)